Recurrent Neural Networks – Models, Capacities, and Applications

Luc De Raedt, Katholieke Universiteit Leuven, Belgium
Barbara Hammer, TU Clausthal, Germany
Pascal Hitzler, Universität Karlsruhe (TH), Germany
Wolfgang Maass, TU Graz, Austria

20.01.08 - 25.01.08

Abstract

The seminar centered around recurrent information processing in neural systems and its connections to brain sciences, on the one hand, and higher symbolic reasoning, on the other side. The goal was to explore connections across the disciplines and to tackle important questions which arise in all sub-disciplines such as representation of temporal information, generalization ability, inference, and learning.

1 Goals of the seminar

Artificial neural networks (FNNs) constitute a particularly successful machine learning technique with application areas ranging from industrial tasks up to simulations of biological neural networks. The seminar centered around recurrent neural networks (RNNs), which include cyclic connections of the neurons and which can incorporate context information or temporal dependencies in a natural way. Spatiotemporal data and context relations occur frequently in various highly relevant domains such as robotics, system identification and control, bioinformatics, medical and biomedical data, sensor streams in technical applications, natural speech processing, analysis of text and web documents, etc. Moreover, spatiotemporal signals and feedback connections are ubiquitous when considering biological neural networks of the human brain. Therefore, RNNs carry the promise of efficient biologically plausible signal processing models optimally suited for a wide area of industrial applications on the one hand and an explanation of cognitive phenomena of the human brain on the other hand.

However, simple feedforward networks without recurrent connections and with a feature encoding of complex spatiotemporal signals which neglects structural

Dagstuhl Seminar Proceedings 08041
Recurrent Neural Networks- Models, Capacities, and Applications
http://drops.dagstuhl.de/opus/volltexte/2008/1424

aspects of data are still the preferred model in industrial or scientific applications, disregarding the great potential of feedback connections. This is mainly due to the fact that traditional training of RNNs, unlike FNNs and backpropagation, faces severe problems: backpropagation for RNNs suffers from numerical barriers, a formal learning theory of RNNs in the classical sense of PAC learning does hardly exist, RNNs easily show complex chaotic behavior which is complicated to manage, and the way how humans use recurrence to cope with language, complex symbols, or logical inference is only partially understood.

The aim of the seminar was to bring together researchers who are involved in these different areas, in order to further the understanding and development of efficient, biologically plausible recurrent information processing, both in theory and in applications. Although often tackled separately, these aspects, the investigation of cognitive models, the design of efficient training models, and the integration of symbolic systems into RNNs, severely influence each other, and they need to be integrated to achieve optimum models, algorithmic design, and theoretical background. Particular aspects which should be addressed in the seminar, included

1. the explanation of 'recurrent' phenomena observed in humans and the development of corresponding biologically plausible models, (*Cognitive models*)
2. the design of efficient recurrent training algorithms beyond numerically instable gradient based techniques, (*Training models*)
3. the theoretical understanding of the capacity of recurrent models, its connections to high-level structures and symbolic paradigms, and the design of corresponding systems. (*Symbolic models*)

2 Structure

29 experts from 10 different countries joined the seminar, including a good mixture of established scientists and promising young researchers working in the field. According to the interdisciplinary topic, the main subjects of the researchers covered heterogeneous fields including computational neuroscience, pattern recognition / neuroinformatics, and logic / relational learning. In spite of the diverse backgrounds, the participants shared a common strong interest in recurrent neural networks, in particular on inference in recurrent neural networks. This unusual setup allowed us to discuss salient issues in a way that integrated perspectives from several scientific disciplines, thereby providing numerous valuable new insights and research contacts for the participants. Correspondingly, a wide range of topics was covered during discussions and brainstorming in the seminar.

During the week, 28 talks were presented which addressed different aspects of RNNs and which were grouped according to the following topics:

- The relation of logic and RNNs

- Statistical logic
- Biological background, experiments and models
- RNN training models
- RNNs for structures

The talks were supplemented by vivid discussions based on the presented topics and beyond. Dedicated discussion sessions centered around various problems and perspectives in this field such as the role of logical inference in RNNs, similarities and dissimilarities of neuro-symbolic integration and statistical relational learning, the parallels between natural and artificial NNs, and current challenges for (industrial) applications of RNNs. This was complemented by slots designated to summarize the insights gained during the week and to put it into a number of questions / challenges. The Wednesday afternoon session 'Trajectory in the environment' in form of a walk in the beautiful surrounding of Dagstuhl gave ample opportunity to further scientific discussions.

3 Results

A variety of open problems and challenges came up during the week. The following topics were identified as central issues in the context of RNNs:

- **Information representation:** How are temporal correlations best represented in RNNs? What is the role of fixed points / attractors / chaotic dynamics in information representation? How can higher symbolic information be represented in RNNs? The problem of how information is represented is central in practical learning tasks as well as biological modelling.
- **RNNs and structures:** How can structures such as sets, graphs, Herbrand-domains be dealt with in RNNs? How can operations on structures, in particular standard planning, logical inference, etc. be realized in (or combined with) RNNs? While a uniform integration of logical structures into RNNs is central in core neuro-symbolic integration, it is tackled in a variety of different ways in concrete learning tasks where structures are involved, ranging from (efficient) ad hoc solutions, dedicated dynamics such as graph NNs, to a tight and efficient marriage of logical representations and statistical modelling in statistical relational learning.
- **Do we need structures:** The question was posed whether higher structures play a role for modelling biological phenomena, on the one hand, and whether structures are beneficial for practical applications, on the other hand. We arrived at the conclusion that higher level structures for inference are likely to be needed also in the brain, but that there are rather sparse biological

data that could suggest suitable models. This gives rise to interesting open questions for neural network researchers, where the challenge is to design artificial neural network models that provide the desired functionality, but also meet known biological constraints. When considering applications, benchmarks such as long-term-prediction, planning, and reasoning inherently rely on structures. It was pointed out that structural aspects can also add a benefit to propositional tasks since problems connected to complex preprocessing of data or bad generalization in the absence of sufficient examples can be avoided by integrating structural aspects.

- **Benchmarks:** One central issue raised in the seminar is the necessity for benchmarks to test the advantages of logic integration and recurrence compared to simple models. Although a number of interesting applications e.g. in chemistry or time series processing exist, large-scale applications where higher-order structures are the key to a good solution in comparison to simpler models seem missing. Thereby, benchmarks can include both, large-scale vectorial domains where the integration of logic or structures yields better / more efficient solutions, as well as large-scale applications from inherently structured domains.

Overall, the presentations and discussions revealed that RNNs constitute a highly diverse and evolving field which opens interesting perspectives to machine learning for structures in the broadest sense. It still waits with quite a few open problems for researchers, a central problem being efficient and robust learning methods for challenging domains.